

ERROR CORRECTION TRELLIS CODING WITH PERIODICALLY INSERTED KNOWN SYMBOLS

Field of the Invention

This invention relates to channel coding, and more particularly, this invention relates to channel coding and trellis codes.

5

Background of the Invention

It is well known that digital data transmitted over a communications channels is subject to signal impairments created by noise, distortion and fading, creating transmission errors. These types of errors can also occur when digital data is stored on devices, including magnetic media or solid-state memories that contain imperfections. This noise is critical in wireless communications where the noise can be extensive and the amount of power used in communication is limited. Typically, the noisy channel is characterized by the signal-to-noise ratio (SNR), which is a ratio of the signal energy over the noise energy in decibels. It is also possible to take the difference in transmission rates into account and define the SNR per bit when the efficiency of codes is compared with different code rates.

It is known that introducing a controlled amount of redundancy by coding in a selected message before a transmission over a communications channel corrects possible errors due to noise, as noted by Dr. Claude Shannon. In a noisy channel, the transmission

rate is reduced to obtain accuracy. A limit, called the Shannon limit, shows the upper bound of the transmission rate for lossless communications in a noisy channel.

5 Block codes were first used in communications, but these block codes were not efficient when comparing to the Shannon limit. More efficient channel codes have been developed on the basis of trellis coding. A trellis encoder includes a
10 Finite State Machine (FSM) and typically a signal mapper. The trellis is generated by the FSM, while code words are binary sequences that are one-to-one mappings of all distinct paths on the trellis by the signal mapper. When a special input sequence is
15 transmitted, the path corresponding to this input sequence is found on the trellis, and the code word generated by the signal mapper. On the decoding side, a maximum likelihood (ML) decoder as a Viterbi decoder or a Maximum a Posteriori (MAP) decoder converts the
20 received code words into the input sequence.

 One important parameter of the Finite State Machine is the memory length m that decides the trellis. The most widely used trellis codes include convolutional codes and turbo codes, as noted in many
25 teachings, including the IEEE technical paper entitled, "Modulation Encoding For Linear Gaussian Channels," IEEE Transactions on Information Theory, Volume 44, No. 6, October 1998, the disclosure which is hereby incorporated by reference in its entirety.

30 Other examples of trellis coding, the use of convolutional codes, and convolutional encoders are disclosed in U.S. Patent Nos. 5,706,312; 6,108,374; 6,215,827 assigned to Lucent Technologies; and 6,347,122 assigned to Agere Systems, the disclosures

which are hereby incorporated by reference in their entirety.

As is known to those skilled in the art, trellis coded modulation is a combined method for joint
5 coding and modulation based on the design of convolutional codes. These are matched to a modulation signal, which is established to maximize the Euclidian distance between a modulated sequence. A trellis diagram is a graph that represents the changing time of
10 a Finite State Machine by defining the states as vertices. Any possible state transitions are displayed as edges. As to the various types of codes, a convolutional code is generated by a linear shift-register circuit and performs a convolution
15 operation on an information sequence, while a turbo code is generated by a parallel concatenation of two (or more) simple convolutional encoders that are separated by interleavers. Thus, along the lines of a turbo code, a concatenated code is established
20 cascading two codes separated by an interleaver. In a generator matrix, a $K \times N$ matrix produces a code word (of length N) of a block code by multiplying the information block of length K . The hamming distance is the number of places in which two sequences of equal
25 length differ. The minimum distance of a code is the smallest hamming distance between any pair of code words. The weight distribution is a list of the hamming distances of each code word from a given reference code word.

30 A Viterbi algorithm is an efficient method for decoding convolutional codes based on finding the path on a trellis diagram that is at a minimum distance from a receive sequence. Related to Viterbi coding is Maximum-Likelihood (ML) decoding that maximizes the
35 probability of a receive sequence given any of the

possible code words. If code words are equally likely, there could be minimum possible error probability.

A convolutional code has been used in GSM and IS-95 systems. Its signal mapper consists of code word
5 generators that define the addition and multiplication on the finite algebra. It is not efficient in comparison of its results to the Shannon limit. Its efficiency would be increased as the memory m increases. Both the computational complexity and
10 memory storage for the Viterbi decoder will be doubled as m increases one. Thus, this approach is not feasible even for $m=20$. Turbo code, as noted before, is a further development of the convolutional code using parallel concatenation of two or more of the
15 convolutional encoders and a pseudo-random interleaver. This code system allows a user to transfer the information in a noisy channel almost without distortion by a rate close to the Shannon limit. Unfortunately, it has a longer delay than that delay
20 used with the conventional convolutional code. In any event, both convolutional codes and turbo codes are recommended in the third generation wireless communication systems, similar to UMTS and CDMA 2000.

25 Summary of the Invention

The present invention overcomes the disadvantages of the prior art indicated above. In the present invention, known symbols, such as zeros, are inserted into a digital input data sequence to form an
30 expanded digital input data sequence. Ordinary trellis coding methods are then applied to the expanded sequence.

In accordance with one aspect of the present invention, a system for channel coding data within a
35 digital communication system includes a data receiving

circuit for receiving a digital input data sequence. This circuit periodically inserts known symbols into the digital input data sequence and forms an expanded digital input data sequence. An encoder is operatively
5 connected to the data receiving circuit for trellis encoding the expanded digital input data sequence and producing a channel coded data stream, such that the number of connections between trellis nodes in a trellis are reduced.

10 In another aspect of the present invention, the known symbols that are inserted comprise zeros. The inserted zeros form a "zero code" that is similar to an equivalent time variant convolutional code.

In yet another aspect of the present
15 invention, the encoder comprises a convolutional encoder and can generate code words that are one-to-one mappings of the distinct paths on a trellis to binary sequences. The topology of the trellis corresponds to memory length m and the known symbols are inserted
20 after each m symbol within the input data sequence.

In yet another aspect of the present invention, the encoder is operative as a generator matrix having a constraint length $K=m-1$, wherein m corresponds to the memory length, and the code rate is
25 $R=1/l$, such that the known symbols are inserted after each $K-1$ information bit. A maximum likelihood (ML) decoder receives and encodes the channel coded data stream. The maximum likelihood (ML) decoder comprises a Viterbi decoder in yet another aspect of the present
30 invention.

A method of channel coding is also disclosed by receiving a digital input data sequence and periodically inserting known symbols into the digital input data sequence and forming an expanded digital
35 input data sequence. The method further comprises the

step of trellis encoding the expanded digital input data sequence to produce a channel coded data stream such that the number of connections between trellis nodes in a trellis are reduced.

5

Brief Description of the Drawings

Other objects, features and advantages of the present invention will become apparent from the detailed description of the invention which follows, when considered in light of the accompanying drawings in which:

10

FIG. 1 is a block diagram showing a basic convolutional encoder that is operative with a data receiving circuit and Viterbi decoder.

15

FIG. 2 illustrates a trellis for an ordinary convolutional code.

FIG. 3 illustrates a trellis where zeros or other known symbols are inserted at the first position periodically.

20

FIG. 4 illustrates a trellis where known symbols, such as zeros, are inserted at the second position periodically.

FIG. 5 illustrates in detail the first calculations used in forming a generator matrix of the present invention.

25

FIG. 6 illustrates a presentation of the generator matrix for a zero code of the present invention, which is similar to a time varying code.

FIG. 6A is a diagram showing first recursive calculations for the decoding algorithm.

30

FIGS. 7A, 7B and 7C illustrate basic steps used in the decoding algorithm of the present invention.

FIG. 8 is a table summarizing properties of the decoding algorithm called the convolutional zero

35

code and comparing it to an ordinary convolutional code.

Detailed Description of the Preferred Embodiments

5 The present invention will now be described more fully hereinafter with reference to the accompanying drawings, in which preferred embodiments of the invention are shown. This invention may, however, be embodied in many different forms and should
10 not be construed as limited to the embodiments set forth herein. Rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art. Like numbers refer to like
15 elements throughout.

 The present invention advantageously is a novel and unobvious channel coding system and method, cooperating with trellis codes, by periodically inserting known symbols, such as zeros, into the
20 transmitted digital data information sequence. By this method and system, the present invention obtains the following advantages: (i) it reduces the computational complexity by half; (ii) it reduces the required memory storages by half; (iii) it reduces the Bit Error Rate
25 (BER). While the delay in the decoding process remains the same, the code rate is reduced by a fraction of the original. In contrast to existing trellis codes, the present invention simplifies the topology of the trellis, i.e., by reducing the number of connections
30 between the nodes in the trellis.

 The following description will progress relative to its implementation on the convolutional code. Naturally, the present invention can be used on many different types of codes as suggested by those
35 skilled in the art. The distance spectrum of a

convolutional code can be improved by the present invention. Simulations have shown that the code rate reduction is compensated by improving the error correction capability of the code.

5 As is known to those skilled in the art, channel coding reduces the Bit Error Rate (BER) by adding redundant bits in the transmitted message. It improves the reception quality by sacrificing the raw data transmission speed. For wireless communications,
10 this is advantageous because limited power is used to transmit information in a noisy channel. It is usually not possible to increase Signal-to-Noise Ratio (SNR) = $10 \log_{10} E_c / 2\sigma^2$ higher than certain level, where E_c is the power of the signals while σ^2 is the variance of noise.
15 On the other hand, if it is possible to work at a low SNR while keeping the same reception quality, then more subscribers are served in a region, provided there is enough bandwidth. The reception quality is measured by Bit Error Rate (BER). Apart from the two properties of
20 (1) BER and (2) the code rate, however, three other properties are also important for wireless communications. These properties are (3) the decoding delay, (4) the computational complexity, and (5) the memory requirement. When designing any mobile
25 communication equipments with limited power and size, all five properties are crucial.

Trellis codes are among the most popular channel codes, and include convolutional codes and turbo codes. Both convolutional and turbo codes have
30 advantages in some of the five areas identified above, but none of them are superior in all five areas. Thus, the development of new channel coding methods that give a good balance among: (1) code rate, (2) BER, (3) delay, (4) computational complexity and (5) memory
35 requirements is difficult.

The present invention is a new channel coding system and method that is operable by periodically inserting known symbols into the transmitted information sequence and then trellis encoding. In one
5 aspect of the present invention, the periodically inserted known symbols are zero symbols, and referred to hereinafter also as the "zero code."

It is well-known by those skilled in the art that the trellis code can be described by a state
10 transfer diagram in a Markov chain (Finite State Machine) and the hamming weights of codewords associated with each transition in the diagram. Efforts have been made to allocate the Hamming weights on the trellis through selecting matrix generators to
15 improve the distance spectrum of the codes. In contrast, by using the inserted known symbols of the present invention, the topology of the trellis is simplified by reducing the number of connections between the trellis nodes. Since both computational
20 complexity and the required memory storages of a Maximum Likelihood (ML) decoder, e.g., the Viterbi decoder, or a Maximum A Posterior (MAP) decoder, depend on the complexity of the trellis, the present invention obtains the benefits on both computation and storage.

25 As a result of inserting zeros, some paths on the trellis will disappear. It is known that the components of the distance spectrum are obtained by searching all possible paths, which start from and end at the zero state and do not pass the zero state, with
30 a certain Hamming weight of the codewords associated to these paths. Thus, at least inserting zeros will not increase the components in the distance spectrum because of the less possible paths in the trellis. There are significant improvements in the distance

spectrum after inserting zeros, and thus, the BER is reduced.

The following description follows the implementation of the present invention on the convolutional code. Naturally, the present invention is not limited to convolutional codes, but can be applied to turbo codes. It is also possible to revive some catastrophic codes into good codes as explained in detail in commonly assigned copending patent application entitled, **"SYSTEM AND METHOD FOR REVIVING CATASTROPHIC CODES,"** filed on the same date, the disclosure which is hereby incorporated by reference in its entirety. The present invention is also operable with a generator matrix such that the zero code of the present invention is equivalent to a special time varying convolutional code. The distance spectrum for some convolutional codes is also improved by the present invention. The corresponding Maximum-Likelihood (ML) decoding algorithm is desired by modifying the Viterbi algorithm.

In trellis coding, different code words are mappings of distinct paths in the trellis through the code generators. To correct possible errors in the received code words, the hamming distances between distinct code words are as large as possible. Usually the optimal encoders are obtained by searching code generators.

Because the code words are associated with the paths in the trellis, they also depend on the topology of the trellis. The topology of a trellis code is simplified dramatically by inserting known symbols, such as zeros, in the present invention. For a convolutional code, the generator matrices for this zero code are found. It will be evident that the

relation and difference between the zero code and the time varying code can be substantial.

In a linear code, it is usually assumed that the information sequence is all-zero while the error bit is one. For the zero code, it is usually not possible to know where the error bit is located, but it is possible to know that it cannot be at the position of the inserted zeros. Thus, the present invention considers (K-1) cases corresponding to different positions of the inserted zero against the error bit one, i.e., there are (K-1) trellis diagrams and (K-1) generator matrices for the zero code.

The present invention uses the time invariant convolutional code with the polynomial generators, with the constraint length $K=m+1$, where m is the memory length, and the code rate is $R=1/l$. It is possible to assume that x_t is binary and the information sequence shifts one bit per time. If $X_k^j = [x_k, x_{k+1}, \dots, x_j]'$, then the code word $c_t = [c_{t,1}, \dots, c_{t,l}]$ and can be identified as:

$$c_{t,i} = g_i X_{t-K+1}^t \pmod{2}, \quad i = 1, 2, \dots, l,$$

and $g_i = [g_i^1, 0, g_i^1, 1, \dots, g_{i,m}^1]$ is the polynomial generator. As usual, it is possible to assume that $x_t=0$ for all $t<0$.

As an example, the optimal code, in the sense that its minimum free distance is the largest, with $K=3$ and $l=2$, is shown in FIG. 1, while its trellis diagram is shown in FIG. 2.

FIG. 1 illustrates the system 10 of channel encoding and decoding that can be used in the present invention where a receiving circuit 12 is illustrated that could be modified to insert known symbols, such as zeros. The encoder 14 is illustrated as a convolutional encoder, but could be other type of encoders as known to those skilled in the art. A

Viterbi decoder 16 is illustrated. The present invention uses a convolutional encoder and a receiving circuit. The trellis diagram for a conventional convolutional encoder where $K=3$ and $L=2$ is shown in

5 FIG. 2.

The dotted lines are indicative for inputting a one, while a darkened line is indicative for inputting a zero. In general, for a trellis with 2^m states, there are 2^{m+1} branches between two consecutive
10 layers.

The present invention inserts one zero after each $(K-1)$ information bits. Since there is always a zero bit in the shift register and its position known, the topology of the convolutional code is changed.

15 Half of the connections between nodes in the trellis are cut off, as shown in the trellis of FIGS. 3 and 4.

In contrast to the original trellis with 2^m states, there are 2^{m-1} states in most layers because there is always a zero in the shift register and its
20 position known. The only exception is when the inserted zero moves to the end of the register. In any case, there are 2^m branches, which are only half of the original, between two consecutive layers after inserting zeros.

25 FIG. 5 illustrates the basic aspects for the generator matrix, showing how the final generator matrix G_j is initiated to result in having its own distance spectrum.

Since $x_{tK+j} = 0$, I can be the information
30 sequence without inserting zeros, and thus:

$$C^{(j)} = IG_j,$$

where G_j is obtained from G by deleting all $(tK+j)$ -th
35 rows, $t = 0, 1, 2, \dots$, in G .

Because G_j is no longer a block Toeplitz matrix, the "zero code" of the present invention is similar to a time varying code. The G_j presentation is shown in the table of FIG. 6, which results because the non-zero sub-block g is not always on the diagonal. Nevertheless, each generator matrix G_j has its own distance spectrum, producing $(K-1)$ distance spectrums for a zero code.

It is also possible to calculate the transfer functions of the convolutional zero codes and compare them to that of the original convolutional code. The results are set forth below.

If $T(w)$ is the transfer function, the ordinary convolutional code with $K=3$, $R=1/2$ and $g_1 = [1, 0, 1]$, $g_2 = [1, 1, 1]$ results in:

$$T(w) = \frac{w^5}{1 - 2w} = w^5 + 2w^6 + \dots + 2^n w^n + \dots$$

In contrast, for the corresponding zero code, no matter whether the first zero is inserted at the first or the second position after the first error bit one results in:

$$T(w) = \frac{w^5}{1 - w} = w^5 + w^6 + \dots + w^n + \dots$$

It is evident that there are improvements on the distance spectrum components at high Hamming distances.

Calculating the transfer function can be difficult for large K . According to the simplified topology of the trellis after periodically inserting zeros, it is possible to calculate the first few components in the distance spectrum for convolutional zero codes. As an example only, some results are

listed to show the improvements in the distance spectrum. When $K=4$, $g_1=17$, $g_2=13$, there is one ordinary spectrum and three zero code spectrums. In the following table, the first row is the hamming distances starting from d_{free} , the second row is the corresponding distance spectrum components for the ordinary convolutional code, and the third, fourth and fifth rows are the distance spectrum components for the convolutional code inserting zeros at the first, second and third positions after the first error bit one, respectively.

	6	7	8	9	10	11	12	13	14	15
	1	3	5	25	25	55	121	267	589	1299
15	0	1	1	3	3	5	8	14	23	39
	1	2	2	7	7	11	19	31	52	85
	1	2	3	8	8	13	21	35	58	97

When $K = 9$, $g_1 = 751$, $g_2 = 557$, it is possible to have one ordinary spectrum and eight zero code spectrums:

	12	13	14	15	16	17	18	19	20
	10	9	30	51	156	340	875	1951	5127
25	5	4	12	21	42	84	180	406	907
	4	5	11	12	36	76	140	305	765
	6	4	18	18	60	125	241	489	1216
	6	4	11	17	46	103	211	427	966
	3	8	11	24	53	95	204	478	972
30	7	9	14	26	64	129	270	571	1317
	10	8	8	15	61	122	266	541	1263
	10	8	17	26	71	142	327	677	1560

The present invention also provides a decoding algorithm. The ML Viterbi decoder searches

the path with the maximal metric among all possible paths on the trellis. Since inserting zeros simplifies the topology of a trellis code, the benefits in three aspects are obtained: computation, storage and BER.

5 The system can append m dummy zeros at the end of the information sequence (zero-tail method). It is also possible to adopt a fixed delay, e.g., a delay $D=5K$, in the Vertibi algorithm.

In the algorithm, $T + \left\lceil \frac{T}{K-1} \right\rceil$ is the total length after

10 inserting zeros, while $[y]$ denotes the integer part of a real number y . Thus, in accordance with the present invention, $x_t = 0$, $t = jK-K-1$ or $t > n$ (formula 2).

$M(X^t_0)$ is the cumulative metric on the path X^t_0 while $L(X^t_{t-K+1})$ is the branch metric associated with the
15 branch X^t_{t-K+1} . The searching $\max_{x_0}^{T-1} M(X^{T-1}_0)$ over all possible paths X^{T-1}_0 can be done by recursive calculating as shown by the algorithm set forth in FIG 6A (equation 3).

 It is known, however, that $X_{t-K+1} = 0$ is the
20 inserted symbol and can be simplified as a formula 4:

$$\phi_t(X^t_{t-K+2}) = L(X^t_{t-K+2}, 0) + \phi_{t-1}(X^{t-1}_{t-K+2}, 0)$$

 There now follows a flow chart shown in FIGS.
25 7A-7C, depicting the various steps of the algorithm. For simplicity in the following charts, the subscript t is dropped and $\phi(X^{t+k-1}_{t+1}) = \phi_t(X^{t+k-1}_{t+1})$ is used. The five step algorithm as shown in FIGS. 7A-7C is obtained.

30 In FIGS. 7A through 7C, a step one initialization is shown (block 30). A step two occurs as a recursive forward algorithm at T (block 32). Step 3 (block 34) indicates a recursive backward algorithm

at T. If $T=N$, then the algorithm proceeds or otherwise reverts back to the recursive forward algorithm (block 36). Termination occurs at step four shown (block 38) and at step five, the recursive backward algorithm is
 5 at the end (block 40).

The properties of the decoding algorithm, i.e., the convolutional "zero code" are compared to an ordinary convolutional code in the table shown in FIG. 8.

10 Simulations for the Additive White Gaussian Noise (AWGN) channel have been accomplished in accordance with the present invention and two examples are illustrated. One example compares the convolutional zero code to the corresponding
 15 convolutional zero code under the same $\text{SNR} = 10 \log_{10} E_c/2\sigma^2$ and the other compares them under the same SNR per bit defined as E_b/N_0 in dB, where $N_0 = 2\sigma^2$ while $E_b = E_c/R$. The simulation was done for $K=91$, $g_1=753$, $g_2=561$ in octal.

20

SNR	-3	-2.5	-2	-1.5	-1
Conv. ($\times 10^{-2}$)	18.8	10.2	4.4	1.4	.34
Zero ($\times 10^{-2}$)	7.3	3.2	1.1	.33	.076
SNR	-0.5	0			
25 Conv. ($\times 10^{-3}$)	.68	.11			
Zero ($\times 10^{-3}$)	.15	.025			
E_b / N_0	0	.5	1	1.5	2
Conv. ($\times 10^{-2}$)	19.0	10.3	4.5	1.5	.35
Zero ($\times 10^{-2}$)	13.9	7.8	3.3	1.3	.32
30 E_b / N_0	2.5	3	3.5	4	4.5
Conv. ($\times 10^{-3}$)	.70	.12			
Zero ($\times 10^{-3}$)	.93	.15			

The first table illustrates that under the same channels, the zero code is superior than the ordinary code. It confirms the improvements in the distance spectrum. The improvement is obtained in the cost of a lower code rate. From the second table, the change in the code rate is taken into account and the zero code is superior at low SNR per bit. The zero code and its corresponding convolutional code usually have the same d_{free} , but fewer components for high hamming weights. The code rate reduction is compensated by the gain in BER.

It is evident that the present invention is an advantageous improvement for trellis coding by periodically inserting known symbols, such as zeros, into the information sequence. This method simplifies the topology of the trellis by reducing the connections between the nodes. Applying it to the convolutional code, it is possible to improve the distance spectrum. While sacrificing the code rate by a fraction, the present invention obtains benefits in computational complexity, storage and BER.

This application is related to copending patent applications entitled, "**SYSTEM AND METHOD FOR REVIVING CATASTROPHIC CODES,**" which is filed on the same date and by the same assignee and inventors, the disclosure which is hereby incorporated by reference.

Many modifications and other embodiments of the invention will come to the mind of one skilled in the art having the benefit of the teachings presented in the foregoing descriptions and the associated drawings. Therefore, it is to be understood that the invention is not to be limited to the specific embodiments disclosed, and that the modifications and embodiments are intended to be included within the scope of the dependent claims.